

Echoes in correlated neural systems

M Helias¹, T Tetzlaff¹ and M Diesmann^{1,2,3}

¹Institute of Neuroscience and Medicine (INM-6),
Computational and Systems Neuroscience, Research
Center Jülich, Germany

²RIKEN Brain Science Institute, Wako, Saitama, Japan

³Medical Faculty, RWTH Aachen University, Germany

E-mail: m.helias@fz-juelich.de

Abstract. Correlations are employed in modern physics to explain microscopic and macroscopic phenomena, like the fractional quantum Hall effect and the Mott insulator state in high temperature superconductors and ultracold atoms. Simultaneously probed neurons in the intact brain reveal correlations between their activity and theoretical work illuminates the importance of correlation for information processing and for macroscopic measures of neural activity, like the EEG. Nevertheless networks of spiking neurons differ from most physical systems: The interaction between elements is directed, time delayed, is mediated by short pulses, and each neuron receives events from thousands of neurons and sends events to thousands of others. Even in the stationary state the network does not reach equilibrium in the sense of detailed balance. Here we develop a quantitative theory of pairwise correlations in finite sized random networks of spiking neurons. We show why the intuitive mean field description fails, how single action potentials reverberate in the network causing an apparent lag of inhibition with respect to excitation, and how global collective oscillations arise.

Keywords: spiking neural networks, correlations, non-equilibrium dynamics, integrate-and-fire model

PACS numbers: 87.18.Sn, 87.19.lc, 87.19.lj, 87.19.lm, 05.10.Gg, 05.70.Ln, 82.40.Bj

AMS classification scheme numbers:

Submitted to: *New J. Phys.*

1. Introduction

Correlations are an established feature of neural activity [1] and evidence increases that they are essential for information processing in the brain [2]. The temporal relationship between the activity of pairs of neurons is described by correlation functions. Their shape has early been related to the direct coupling between neurons and to the common input shared by pairs of neurons. Correlations affect the representation of information as they may limit the signal-to-noise ratio of population rate signals [3]. On the other hand, they have been shown to increase the amount of information available to unbiased observers [4]. Furthermore, synchronous neural activity is thought to bind elementary representations into more complex objects [5] and experimental evidence for such a correlation code is provided by task related modulation of synchrony in primary visual cortex [6] and in motor cortex [7].

The small magnitude [8] of pairwise correlations in the asynchronous irregular state [9] of cortex has recently been related to the balance between excitation and inhibition in local networks [10, 11] and inhibitory feedback was identified as a general mechanism of decorrelation [12]. However, a quantitative theory explaining the temporal shape of correlation functions in recurrently impulse coupled networks of excitatory and inhibitory cells remained elusive.

Assuming random connectivity with identical numbers and strengths of incoming synapses per neuron, as illustrated in Figure 1, suggests by mean field arguments [13, 14] that the resulting activity of two arbitrarily selected neurons and hence the power spectra of activities averaged over excitatory or inhibitory neurons should be the same. Direct simulations, however, exhibit different power spectra for both sub-populations [15]. A similar argument holds for the covariance c_{ff} between the two neurons: If the covariance c between any pair of inputs is known, the covariance between their outgoing activity c_{ff} is fully determined. By self-consistency, as both neurons belong to the same recurrent network, one concludes that $c_{ff} = c$. In particular the covariance averaged over excitatory pairs should be identical to the corresponding average over inhibitory pairs, which is in contrast to direct simulation (Figure 1b). In this work, we elucidate why this mean field argument for covariances fails and derive a self-consistency equation for pairwise covariances in recurrent random networks which explains the differences in the power spectra and covariances.

Theories for pairwise covariances have been derived for binary neuron models

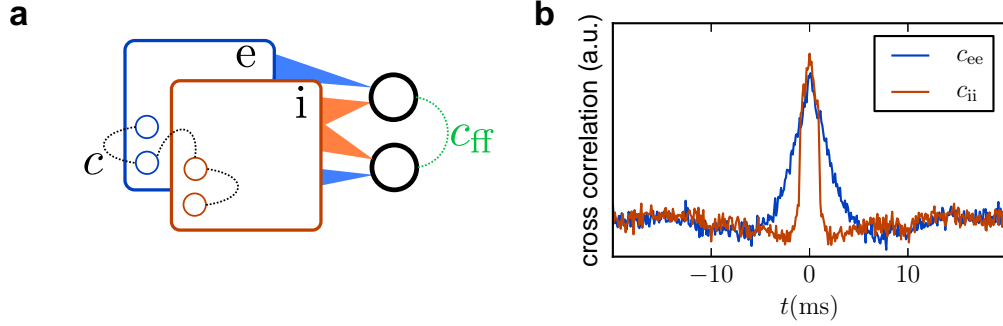


Figure 1. Self-consistency argument fails for covariances in a homogeneous recurrent random network. **(a)** Each neuron (black circles) receives input from the same number of randomly chosen excitatory (e) and inhibitory (i) neurons in the network, so the input statistics of all neurons is the same. The covariance c within the network determines the covariance between the neuron inputs to a pair of neurons and hence the covariance c_{ff} of their outputs. Self-consistency seems to require $c_{ff} = c = c_{ee} = c_{ii}$. **(b)** Covariance functions averaged over pairs of excitatory (c_{ee}) and over pairs of inhibitory (c_{ii}) integrate-and-fire model neurons are different in a direct simulation. The other parameters are given in [Appendix E](#).

[11, 16] and for excitatory stochastic point process models [17]. However, the lack of either inhibition or delayed pulsed interaction limits the explanatory power of these models. A theory for networks of leaky integrate-and-fire model neurons [18] is required, because this model has been shown to well approximate the properties of mammalian pyramidal neurons [19] and novel experimental techniques allow to reliably assess the temporal structure of correlations in cortex [20]. Moreover, the relative timing of action potentials is the basis for models of synaptic plasticity [21], underlying learning in biological neural networks. Analytical methods to treat population fluctuations in spiking networks are well advanced [22] and efficient hybrid analytical-numerical schemes exist to describe pairwise covariances [23]. Here we present an analytically solvable theory of pairwise covariances in random networks of spiking leaky integrate-and-fire model neurons with delayed pulsed interaction in the asynchronous irregular regime.

2. Results

We consider recurrent random networks of N excitatory and γN inhibitory leaky integrate-and-fire model neurons receiving pulsed input (spikes) from other neurons

in the network. Each neuron has $K = pN$ incoming excitatory synapses independently and randomly drawn from the pool of excitatory neurons, and $\gamma K = \gamma pN$ inhibitory synapses (homogeneous Erdős-Rényi random network with fixed in-degree). An impulse at time t arrives at the target neuron after the synaptic delay d and elicits a synaptic current I_i that decays with time constant τ_s and causes a response in the membrane potential V_i (with time constant τ_m) proportional to the synaptic efficacy J (excitatory) or $-gJ$ (inhibitory), respectively. The coupled set of differential equations governing the subthreshold dynamics of a single neuron i is [24]

$$\begin{aligned}\tau_m \frac{dV_i}{dt} &= -V_i + I_i(t) \\ \tau_s \frac{dI_i}{dt} &= -I_i + \tau_m \sum_{j=1, j \neq i}^N J_{ij} s_j(t-d),\end{aligned}\tag{1}$$

where the membrane resistance was absorbed into J_{ij} . If V_i reaches the threshold V_θ at time point t_k^i the neuron emits an action potential and the membrane potential is reset to V_r , where it is clamped for the refractory time τ_r . The spiking activity of neuron i is described by this sequence of action potentials, the spike train $s_i(t) = \sum_k \delta(t - t_k^i)$.

The activity of a given neuron i depends on the history of the other neurons' activities $\mathbf{s}(t) = (s_1(t), \dots, s_N(t))^T$ in the network. The time averaged covariance matrix expresses these interrelations and is defined as $\bar{\mathbf{c}}(\tau) = \langle \mathbf{s}(t+\tau) \mathbf{s}^T(t) \rangle_t - \mathbf{r} \mathbf{r}^T$, where $\mathbf{r} = \langle \mathbf{s} \rangle_t$ is the vector of time averaged firing rates. The diagonal contains the autocovariance functions (diagonal matrix $\mathbf{a}(\tau)$) which are dominated by a δ -peak at zero time lag and for $\tau \neq 0$ exhibit a continuous shape mostly determined by refractoriness, the inability of the neuron to fire spikes in short succession due to the voltage reset, as shown in Figure 2d. The off-diagonal elements contain the cross covariance functions $\mathbf{c}(\tau)$ that originate from interactions. We therefore decompose the covariance matrix into $\bar{\mathbf{c}} = \mathbf{a} + \mathbf{c}$. A basic property of covariance matrices is the symmetry $\bar{\mathbf{c}}(\tau) = \bar{\mathbf{c}}^T(-\tau)$, so we only need to consider $\tau > 0$ and obtain the solution for $\tau < 0$ by symmetry. Each spike at time t' can influence the other neurons at time $t > t'$. Formally we express this influence of the history up to time t on a particular neuron i in the network as $s_i(t, \{\mathbf{s}(t') | t' < t\})$, which is a functional of all spike times until t . In the asynchronous state of the network [9] and for small synaptic amplitudes J_{ij} a single spike of neuron j causes only a small perturbation of s_i . The other inputs

to neuron i effectively act as additional noise. We may therefore perform a linear approximation of j 's direct influence on neuron i and average over the realizations of the remaining inputs $\mathbf{s} \setminus s_j$, as illustrated in [Figure 2a](#). This linearization and the superposition principle lead to the convolution equation

$$\langle s_i(t) | s_j \rangle_{\mathbf{s} \setminus s_j} = r_i + \int_{-\infty}^t h_{ij}(t, t') (s_j(t') - r_j) dt' \quad (2)$$

$$= r_i + [h_{ij} * (s_j - r_j)](t), \quad (3)$$

where we define as the linear response kernel $h_{ij}(t, t') = \left\langle \frac{\delta s_i(t)}{\delta s_j(t')} \right\rangle_{\mathbf{s} \setminus s_j}$ the functional derivative of $s_i(t)$ with respect to $s_j(t')$, formally defined in [Appendix A](#). This kernel quantifies the effect of a single spike at time t' of neuron j on the expected density of spikes of neuron i at time t by a direct synaptic connection from neuron j to neuron i . This density vanishes for $t < t'$ due to causality. For the stationary network state studied here, the kernel further only depends on the time difference $\tau = t - t'$. In a consistent linear approximation there are no higher order terms, so the effects of two inputs s_j and s_k superimpose linearly. Using this linear expansion and the definition of the covariance matrix, the off-diagonal elements of the covariance matrix fulfill a linear convolution equation

$$\mathbf{c}(\tau) = [\mathbf{h} * (\mathbf{a} + \mathbf{c})](\tau) \quad \text{for } \tau > 0. \quad (4)$$

The equation is by construction valid for $\tau > 0$ and needs to be solved simultaneously obeying the symmetry condition $\mathbf{c}(t) = \mathbf{c}(-t)^T$. Up to linear order in the interaction, equation (4) is the correct replacement of the intuitive self-consistency argument sketched in [Figure 1](#).

In order to relate the kernel h_{ij} to the leaky integrate-and-fire model, we employ earlier results based on Fokker-Planck theory [\[24\]](#). For small synaptic amplitudes $J \ll V_\theta - V_r$ and weak pairwise covariances the summed synaptic input $\tau_m \sum_{j=1}^N J_{ij} s_j(t - d)$ can be approximated as a Gaussian white noise with mean $\mu_i = \tau_m \sum_j J_{ij} r_j$ and variance $\sigma_i^2 = \tau_m \sum_j J_{ij}^2 r_j$ for neurons firing with rates r_j and Poisson statistics. For short synaptic time constants $\tau_s \ll \tau_m$ the stationary firing rate $r_i(\mu_i, \sigma_i^2)$ ([A.1](#)) depends on these two moments [\[24\]](#). In a homogeneous random recurrent network the input to each neuron is statistically the same, so the stationary rate $r_i = r$ is identical for all neurons. It is determined by the self-consistent solution of $r(\mu_i, \sigma_i^2)$, taking into account the dependence of μ_i and σ_i^2 on the rate r itself [\[9\]](#).

The integral $w_{ij} = \int_0^\infty h_{ij}(t) dt$ of the response kernel is equivalent to the DC

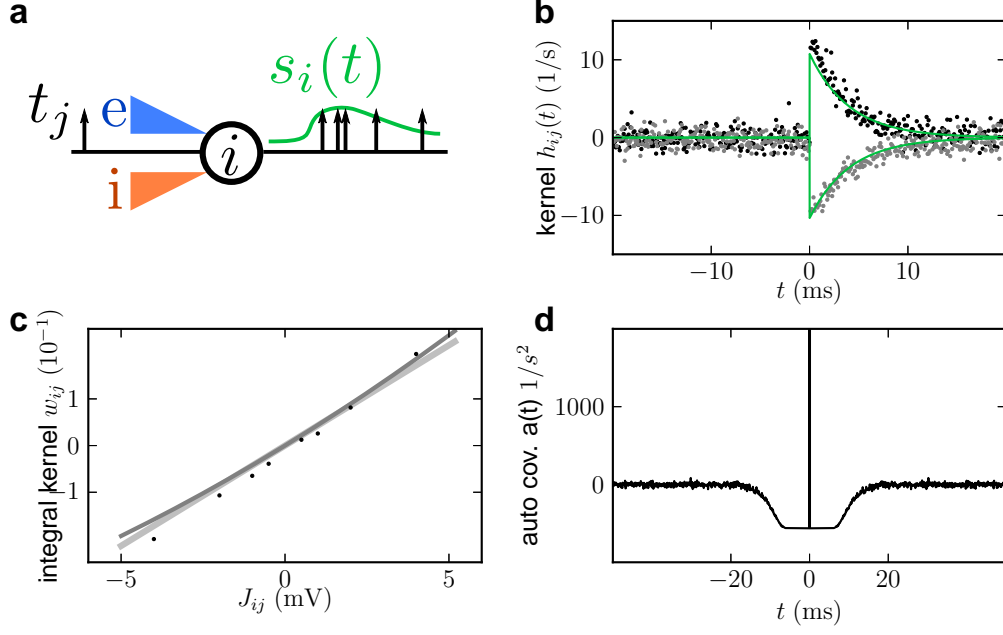


Figure 2. Mapping the integrate-and-fire dynamics to a linear coupling kernel. **(a)** The kernel h_{ij} determines the transient effect of an incoming impulse at time point t_j (black arrow) on the density $s_i(t)$ of outgoing action potentials, averaged over realizations of the stochastic activity of the remaining inputs (indicated as red and blue triangles). **(b)** Response kernel (5) (green) compared to direct simulation for an impulse of amplitude $J = 1$ mV (black dots) and $J = -1$ mV (gray dots). Time constant $\tau_e = 4.07$ ms determined by a least squares fit to a single exponential. The background activity causes a mean $\mu_i = 15$ mV and fluctuations $\sigma_i = 10$ mV. **(c)** Linear and quadratic dependence (A.3) of the integral response w_{ij} on J_{ij} (dark gray curve) and linear term alone (light gray line). **(d)** Autocovariance function of the spike train with a δ peak at $t = 0$ and covariance trough due to refractoriness.

susceptibility $w_{ij} = \frac{\partial r_i}{\partial r_j}$ [25], leading to an approximation of second order in the synaptic amplitude $w_{ij} = \alpha J_{ij} + \beta J_{ij}^2$. The first order term originates from the dependence of μ_i on r_j , the second order term stems from the dependence of σ_i^2 on r_j (A.3).

Figure 2b shows the deflection of the firing rate from baseline caused by an impulse in the input averaged over many repetitions. For sufficiently strong fluctuations σ_i , the quadratic term in J_{ij} is negligible, as seen from Figure 2c. The kernel shows exponential relaxation with an effective time constant τ_e , that depends

on the working point (μ_i, σ_i) and the parameters of the neuron, which is obtained by a least squares fit of a single exponential to the simulated response in Figure 2b for one particular amplitude J_{ij} . We therefore approximate the response $h_{ij}(t)$ as

$$w_{ij}h(t) = \Theta(t - d) \frac{w_{ij}}{\tau_e} e^{-\frac{t-d}{\tau_e}}, \quad (5)$$

where d is the synaptic delay and Θ the Heaviside step function.

In experiments covariance functions are typically averaged over statistically equivalent pairs of neurons. Such averages are important, because they determine the behavior of the network on the macroscopic scale of populations of neurons. We therefore aim at a corresponding effective theory. Assuming identical dynamics (1), all neurons have to good approximation the same autocovariance function $a(t)$ and response kernel $h(t)$. So each incoming excitatory impulse causes a response $w h(t)$, an inhibitory impulse $-g w h(t)$. We define the covariance function averaged over all pairs of excitatory neurons as $c_{ee}(\tau) = \frac{1}{N^2} \sum_{i,j \in \mathcal{E}, i \neq j} c_{ij}(\tau)$, (setting $N(N-1) \simeq N^2$ for $N \gg 1$), where \mathcal{E} denotes the set of all excitatory neurons. The pairings c_{ei} , c_{ie} , and c_{ii} are defined analogously. Inserting equation (4) into the average $c_{ee}(\tau)$, the first term proportional to the autocovariance $a(t)$ only contributes if neuron j projects to neuron i . For fixed i , there are K such indices j , so the first term yields $\sum_{i,j \in \mathcal{E}, i \neq j} h_{ij} * a_j = NK w h * a$. The second sum $\sum_{i,j \in \mathcal{E}, i \neq j} h_{ik} * c_{kj}$ can be decomposed into a sum over all intermediate excitatory neurons $k \in \mathcal{E}$ and over all inhibitory neurons $k \in \mathcal{I}$ projecting to neuron i . Replacing the individual covariances by their population average, c_{ee} and c_{ie} , respectively, and considering the number of connections and their amplitude we obtain $wNK h * (c_{ee} - \gamma g c_{ie})$. Similar relations hold for the remaining three averages, so we arrive at a two-by-two convolution matrix equation for the pairwise averaged covariances for $\tau > 0$

$$\begin{aligned} \mathbf{c}(\tau) &= [h * \mathbf{M} \mathbf{c}] (\tau) + \mathbf{Q} [h * a] (\tau) \\ \text{with } \mathbf{M} &= Kw \begin{pmatrix} 1 & -\gamma g \\ 1 & -\gamma g \end{pmatrix}, \quad \mathbf{Q} = \frac{Kw}{N} \begin{pmatrix} 1 & -g \\ 1 & -g \end{pmatrix}, \\ \text{and } \mathbf{c}(\tau) &= \begin{pmatrix} c_{ee}(\tau) & c_{ei}(\tau) \\ c_{ie}(\tau) & c_{ii}(\tau) \end{pmatrix}. \end{aligned} \quad (6)$$

The convolution equation only holds for positive time lags τ . For negative time lags it is determined by the symmetry $\mathbf{c}(-\tau) = \mathbf{c}^T(\tau)$. The solution

of this equation can be obtained by an extension of the method used in [17] employing Wiener-Hopf theory [26] to the cross spectrum $\mathbf{C}(\omega) = \int_{-\infty}^{\infty} \mathbf{c}(t) e^{-i\omega t} dt$ in frequency domain, as shown in Appendix B (here capital letters denote the Fourier transform of the respective lower case letters). With the definition of the propagator $\mathbf{P}(\omega) = (\mathbf{1} - \mathbf{M}H(\omega))^{-1}$ the cross spectrum takes the form $\mathbf{C}(\omega) = \mathbf{P}(\omega) [D_+(\omega)\mathbf{Q} + D_+(-\omega)\mathbf{Q}^T - A(\omega)|H(\omega)|^2\mathbf{Q}\mathbf{M}^T] \mathbf{P}^T(-\omega)$, where we split the term $H(\omega) A(\omega) = D_+(\omega) + D_-(\omega)$ so that $d_+(\tau)$ and $d_-(\tau)$ vanish for times $\tau < 0$ and $\tau > 0$, respectively. The non-averaged cross spectrum can be recovered as a special case setting $\mathbf{Q} = \mathbf{M} = \mathbf{W}$, because the convolution equations (4) and (6) have the same structure and symmetries. If all eigenvalues of $\mathbf{M}H(\omega)$ have an absolute value smaller unity, the propagator $\mathbf{P}(\omega)$ can be expanded into a geometric series in which the n -th term contains only interactions via n steps in the connectivity graph. This expansion has been used to obtain the contribution of different motifs to the integral of covariance functions [27] and their temporal shape [28].

In the following, we neglect the continuous part of the autocovariance function, setting $a(t) = r \delta(t)$, because the δ -peak is typically dominant. For an arbitrary causal kernel h it follows that $D_+(\omega) = r H(\omega)$, so the cross spectrum takes the form

$$\begin{aligned} \mathbf{C}(\omega) = & r \frac{Kw}{N} \begin{pmatrix} 1 & -g \\ 1 & -g \end{pmatrix} U(i\omega) + \text{c.c. trans.} \\ & + r(1 + g^2\gamma) \frac{(Kw)^2}{N} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} |U(i\omega)|^2 \\ \text{with } U(z) = & \frac{1}{H^{-1}(-iz) - L} \\ \text{and } L = & Kw(1 - \gamma g). \end{aligned} \quad (7)$$

The limit $\omega \rightarrow 0$ corresponds to the time integral of the cross covariance function, approximating the count covariance for long time bins [2, 8]. With $A(0) = r$, the integral correlation coefficient averaged over neuron pairs fulfills the equation

$$\begin{aligned} \frac{\mathbf{C}(0)}{A(0)} = & \frac{Kw}{N} \frac{1}{1-L} \begin{pmatrix} 2 & 1-g \\ 1-g & -2g \end{pmatrix} \\ & + \frac{(Kw)^2}{N} \frac{1+g^2\gamma}{(1-L)^2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \end{aligned} \quad (8)$$

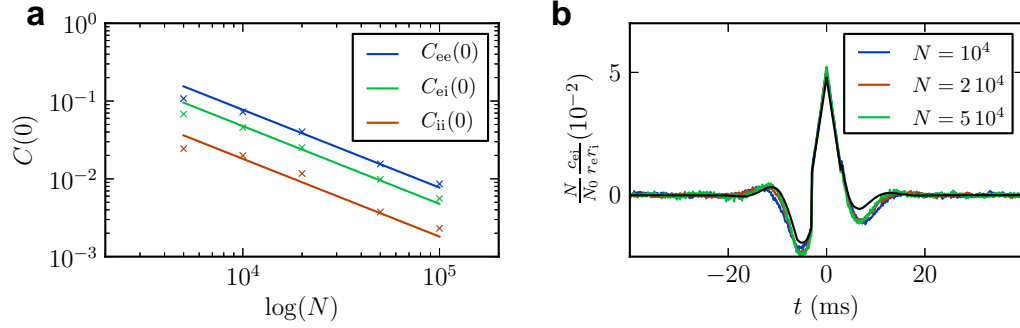


Figure 3. Shape invariance of covariances with network scale. Synapses scaled as $J \propto 1/N$ with network size N to conserve the population feedback $L = \text{const.}$ (a) Integral correlation coefficient averaged over different pairs of neurons and theory (8) confirming the $\propto 1/N$ dependence. (b) Rescaled covariance functions $N C_{ei}$ averaged over excitatory-inhibitory pairs of neurons for different network sizes (color coded) and theory (15) (black).

which has previously been derived from a noise-driven linear rate dynamics [12]. The quantity L plays a key role here: It determines the feedback magnitude of in-phase fluctuations of the excitatory and inhibitory population. Stability of the average firing rate requires this feedback to be sufficiently small [9], i.e. $L < 1$, indicated by the pole in equation (8) at $L = 1$. Typically cortical networks are in the balanced regime [9, 13, 14], i.e. $L < 0$. For such inhibition dominated networks, the denominator in equation (8) is larger than unity, indicating a suppression of covariances [12]. As shown in Figure 3a, the prediction (8) agrees well with the results of simulations of leaky integrate-and-fire networks for a broad range of network sizes N .

Previous works have investigated neural networks in the thermodynamic limit $N \rightarrow \infty$ [11, 13], scaling the synaptic amplitudes to zero in order to arrive at analytical results. Equation (8) provides an alternative criterion to scale the synapses while keeping the dynamics comparable: The two terms in equation (8) depend differently on the feedback L . In order to maintain their ratio, we need to keep the population feedback L constant. The synaptic amplitude J (approximately $\propto w$ (A.3)) hence needs to scale as $J \propto 1/N$. In addition, the response kernel h of each single neuron must remain unchanged, requiring the same working point, characterized by the mean μ_i and fluctuations σ_i in the input into each cell. Constant mean directly follows from $L = \text{const.}$, but the variance due to local input from other

neurons in the network decreases as $1/N$. To compensate, we supply each neuron with an additional external uncorrelated balanced noise whose variance appropriately increases with N (as described in detail in [Appendix E](#)). [Figure 3b](#) shows that the shape of the covariance functions is invariant over a large range of network sizes N , in particular the apparent time lag of inhibition behind excitation observed as the asymmetry in [Figure 3b](#) does not vanish in the limit $N \rightarrow \infty$. The magnitude of the covariance decreases as $1/N$ as expected from equation (8), because $Kw = \text{const.}$ Previous studies have applied a different scaling $J \propto 1/\sqrt{N}$ in order to keep the variance at the input to each neuron constant [11]. Such a scaling increases the feedback on the network level $L \propto \sqrt{N}$ and therefore changes the collective network state.

Global properties of the network dynamics can be inferred by considering the spectrum of equation (7), those complex frequencies z_k at which the expression has a pole due to the function $U(z)$. These poles are resonant modes of the network, where the real part $\Re(z_k)$ denotes the damping of the mode, and the imaginary part $\Im(z_k)$ is the oscillation frequency. A pole appears whenever z_k is a single root of $U^{-1}(z_k) = H^{-1}(-iz_k) - L = 0$. With the Fourier representation $H(\omega) = \frac{e^{-i\omega d}}{1+i\omega\tau_e}$ of the response kernel (5), the poles correspond to the spectrum of the delay differential equation $\tau_e \frac{dy}{dt}(t) = -y(t) + Ly(t-d)$, (cf. [29]) which describes the evolution of the population averaged activity. As shown in [Appendix C](#), the location of the poles can be expressed by the branches k of the Lambert- W function, the solution of $W_k e^{W_k} = x$ [30], as

$$z_k = -\frac{1}{\tau_e} + \frac{1}{d} W_k \left(L \frac{d}{\tau_e} e^{\frac{d}{\tau_e}} \right) \quad k \in \mathbb{N}_0. \quad (9)$$

The spectrum only depends on the population feedback L , the delay d and the effective time constant τ_e of the neural response kernel. This explains why keeping L constant while scaling the network in [Figure 3](#) yields shape invariant covariance functions. A typical spectrum is shown in [Figure 4a](#) as an inset, where each dot marks one of the solutions z_k of equation (9). The two principal branches of W_k are the modes with the largest real part $\Re(z_k)$, and hence with the least damping, dominating the network dynamics. The remaining branches appear as conjugate pairs and their real parts are more negative, corresponding to stronger damping. Investigating the location of the principal branches therefore enables us to classify the dynamics in the network. Their dependence on the delay is shown in [Figure 4a](#) as a parametric plot in d . The point at which the two real principal solutions turn into a complex conjugate

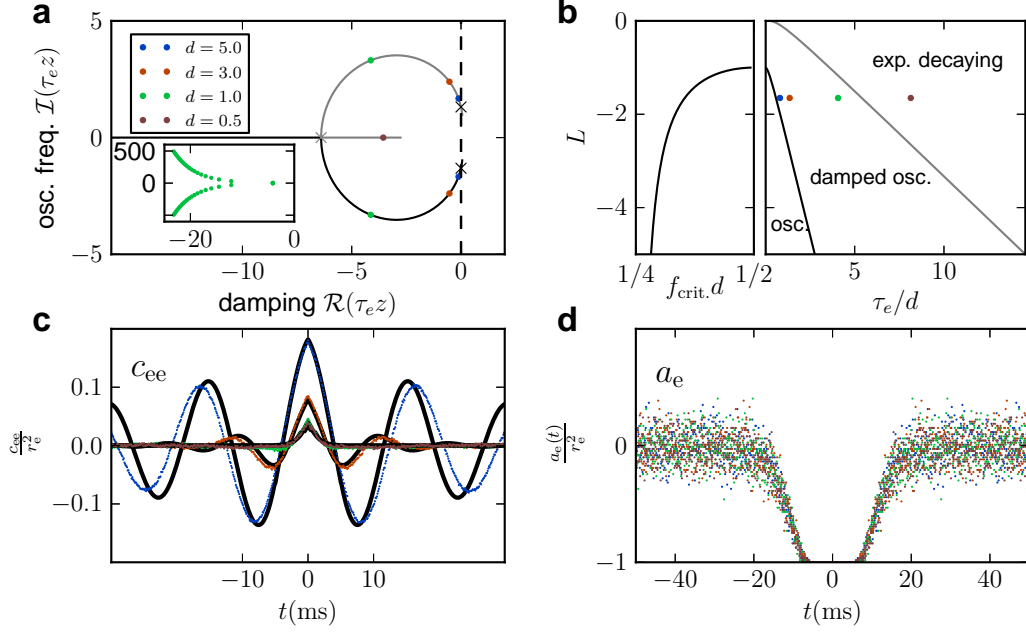


Figure 4. Phase diagram determined by spectral analysis. **(a)** The two rightmost poles change with delay d (colors correspond to delays throughout). At $d = 0.753$ ms (gray cross) (10) the poles become a conjugate pair, at $d = 6.88$ ms (black crosses) (12) both poles have a zero real part, causing oscillations (Hopf bifurcation). Each dot in the inset represents a pole z_k (9) for delay $d = 1$ ms (two rightmost poles appear as one point). **(b)** Right: Phase diagram spanned by τ_e/d and feedback L . Onset of oscillations below the black curve (12) (Hopf bifurcation, black crosses in **a**), damped oscillations below the gray curve (10) (gray cross in **a**). Left: Oscillation frequency (12) at the Hopf bifurcation. **(c)** Average cross covariance between excitatory neurons and theory (15) (black). **(d)** Average autocovariance of excitatory neurons (δ -peak not shown).

pair marks a transition from purely exponentially decaying dynamics to damped oscillations. This happens at sufficiently strong negative coupling L or sufficiently long delay d , precisely when the argument of W_k is smaller than $-e^{-1}$ [30], leading to the condition

$$L < -\frac{\tau_e}{d} e^{-\frac{d}{\tau_e}-1}. \quad (10)$$

The gray cross marks this point in Figure 4a, the gray curve shows the corresponding relation of feedback and delay in the phase diagram Figure 4b. In the region below the curve, the dominant mode of fluctuations in the network is thus damped

oscillatory, whereas above the curve fluctuations are relaxing exponentially in time.

For sufficiently long delay d the principal poles may assume positive real values, leading to ongoing oscillations, a Hopf bifurcation. The condition under which this happens can be derived from $U^{-1}(\omega_{\text{crit.}}) - L = 0$, as detailed in [Appendix C](#). Equating the absolute values on both sides leads to the condition $\omega_{\text{crit.}}\tau_e = \sqrt{L^2 - 1}$: oscillations can only be sustained, if the negative population feedback is sufficiently strong $L < -1$. The oscillation frequency increases the stronger the negative feedback. The condition for the phases leads to the critical delay required for the onset of oscillations (see [Appendix C](#) for details)

$$\frac{d_{\text{crit.}}}{\tau_e} = \frac{\pi - \arctan(\sqrt{L^2 - 1})}{\sqrt{L^2 - 1}}. \quad (11)$$

This relation is shown as the black curve in the phase diagram [Figure 4b](#). The oscillatory frequency on the bifurcation line, at the onset of oscillations can be expressed as

$$2\pi f_{\text{crit.}}d = \omega_{\text{crit.}}d = \pi - \arctan(\sqrt{L^2 - 1}), \quad (12)$$

which is shown in the left sub-panel of the phase diagram [Figure 4b](#). Consequently, the oscillation frequency $f_{\text{crit.}}$ at the onset is between $(4d_{\text{crit.}})^{-1}$ and $(2d_{\text{crit.}})^{-1}$, depending on the strength of the feedback L , with increasing negative feedback approaching $f_{\text{crit.}} = (4d_{\text{crit.}})^{-1}$ at the onset.

Changing the synaptic delay homogeneously for all synapses in the network allows to observe the transition of the network from exponentially damped, to oscillatory damped, and finally to oscillatory dynamics. For a short delay of $d = 0.5$ ms the dynamics is dominated by the single real pole near $-1/\tau_e$ (brown dot in [Figure 4a](#)) and the covariance function is exponentially decaying ([Figure 4c](#)). Increasing the delay to $d = 1$ ms the principal poles split into a complex conjugate pair as the delay crosses the gray curve in [Figure 4b](#) so that side troughs become visible in the covariance function in [Figure 4c](#). Further increasing the delay, the network approaches the point of oscillatory instability, where a Hopf bifurcation occurs, marked by black crosses in [Figure 4a](#) and the black curve in [Figure 4b](#). The damping of oscillations decreases as the system approaches the bifurcation ([Figure 4c](#)). The auto covariance function of single spike trains, however, stays mostly unchanged ([Figure 4d](#)). This state is known as synchronous irregular activity [9], where the activity of a single neuron is irregular, but collectively the neurons participate in a global oscillation.

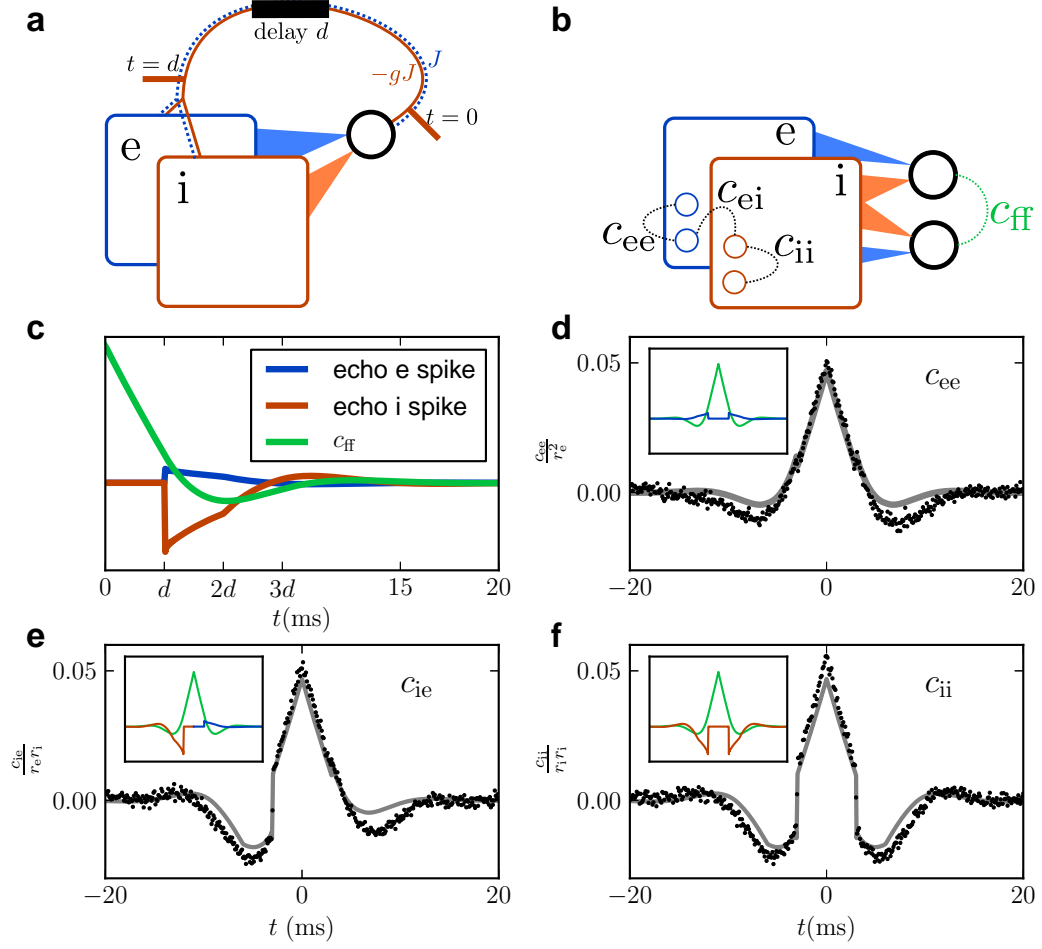


Figure 5. Composition of covariance functions. (a) Network echo caused by a spike sent at time $t = 0$ sets in after one synaptic delay ($d = 3$ ms): red curve in c, inhibitory spike, blue curve in c, excitatory spike. (b) Correlated inputs from the network to a pair of neurons (black circles) cause covariance c_{ff} between their outputs (green curve in c). (d)-(f) Covariance functions averaged over pairs of neurons (black dots, d: excitatory pairs, e: excitatory-inhibitory pairs, f: inhibitory pairs) and theory (15) (gray underlying curve). Inset shows the two components from c that add up.

If the network is not in the oscillatory state, all modes are damped in time, so all poles z_k of the function $U(z)$ appearing in equation (7) lie in the left complex half plane, $\Re(z_k) < 0$; the function $|U(z)|^2$ has poles in both half-planes. We perform the Fourier transform to time domain using the residue theorem $u(t) = \frac{1}{2\pi i} \oint_{E_t} U(z) e^{zt} dz = \sum_{z_k \in E_t} \text{Res}(U, z_k) e^{z_k t}$, where the integration path E_t proceeds along the imaginary axis from $-i\infty$ to $i\infty$ and is then closed in infinity in the left half plane (for $t > d$) or the right half plane (for $0 < t < d$) to ensure convergence, resulting in (see [Appendix D](#) for the detailed calculation)

$$u(t) = \sum_{k=0}^{\infty} \frac{1}{(1 + z_k \tau_e) d + \tau_e} \Theta(t - d) e^{z_k(t-d)}. \quad (13)$$

The back transform of $V(\omega) \stackrel{\text{def}}{=} |U(\omega)|^2$ proceeds along similar lines and results in

$$v(t) = \sum_{k=0}^{\infty} \frac{1}{(1 + z_k \tau_e) d + \tau_e} \frac{e^{z_k|t|}}{(1 - z_k \tau_e) - L e^{z_k d}}. \quad (14)$$

The population averaged covariance functions in the time domain then follows from equation (7) for $t > 0$ as

$$\begin{aligned} \mathbf{c}(t) = & r \frac{Kw}{N} \begin{pmatrix} 1 & -g \\ 1 & -g \end{pmatrix} u(t) \\ & + r \frac{(Kw)^2}{N} (1 + g^2 \gamma) \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} v(t), \end{aligned} \quad (15)$$

which is the central result of our work. [Figure 5](#) shows the comparison of the theory (15) with the covariance functions obtained by direct simulation. The analytical expression unveils that the covariance functions are composed of two components: the first line in equation (15) has heterogeneous matrix elements and hence depends on the neuron types under consideration. Its origin is illustrated in [Figure 5a](#): If one of the neurons emits a spike, as indicated, this impulse travels along its axon and reaches the target neurons after one synaptic delay d . Depending on the type of the source neuron, the impulse excites (synaptic amplitude J) or inhibits ($-gJ$) its targets. Its effect is therefore visible in the pairwise covariance function as a

positive (blue) or negative (red) deflection, respectively in Figure 5c. This deflection not only contains the direct synaptic effect, but also infinitely many reverberations of the network, seen formally in equation (13). This expression is not proportional to the kernel $h(t)$ directly, but is rather a series including the whole spectrum of the network. The shape of the spike echo consequently shows onsets of reverberations at integer multiples of the synaptic delay (Figure 5c), being transmitted over multiple synaptic connections. The contribution of the second line in equation (15) follows the intuitive argument illustrated in Figure 5b. The incoming activity from the network to a pair of neurons is correlated. As the input statistics is the same for each neuron, this contribution is identical for any pair of neurons (green curve in Figure 5c). The sum of both components results in the covariance functions shown in Figure 5d-f. The same analytical solution is shown for different delays in Figure 4c showing good agreement with direct simulation. For different sizes of simulated networks in Figure 3b-d the analytical expression (15) explains why the spike echo does not become negligible in the thermodynamic limit $N \rightarrow \infty$: for fixed population feedback L , both contributions in equation (15) scale as $1/N$, so the relative contribution of the echo stays the same. This also explains the apparent paradox (see Figure 1), that covariance functions in recurrent networks not only depend on the input statistics, but in addition the spike feedback causes a reverberating echo. The power spectrum of population-averages is dominated by pairwise covariances, explaining the different spectra observed in the excitatory and inhibitory population activity [15]. Scaling the network such as to keep the marginal statistics of single neurons constant, $J \propto w \propto 1/\sqrt{N}$ [11, 13] changes the spectrum (9), because the feedback increases as $L \propto \sqrt{N}$ which can ultimately lead to oscillations as shown in Figure 4c.

3. Discussion

The present work qualitatively explains certain features of the correlation structure of simultaneously recorded synaptic currents of two cells in vivo. Novel experimental techniques are able to separate contributions of excitatory and inhibitory inputs [20]. We calculate such covariances in a random network and show that the covariance between synaptic impulses decomposes into a linear combination of the covariance of the spiking activity and the autocovariance functions (see caption of Figure 6). Each synaptic impulse has a certain time course, here modeled as a first order low pass with time constant $\tau_s = 2$ ms (1). The covariance between these filtered currents is

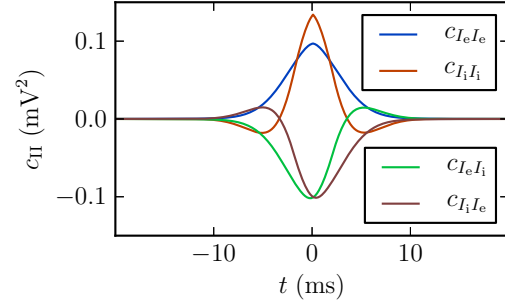


Figure 6. Covariance between synaptic currents of a pair of neurons in a recurrent random network in analogy to in vivo experiments [20]. Covariance of excitatory contributions (blue, $c_{I_e I_e} = q * (J^2 p K a_e + J^2 K^2 c_{ee})$), analogously between inhibitory contributions (red), between excitatory and inhibitory contribution (green, $c_{I_e I_i} = -q * (g J^2 \gamma K^2 c_{ei})$), and $c_{I_i I_e}$ analogously (brown). Currents are filtered by an exponential kernel of time constant $\tau_s = 2$ ms (1), leading to the filtering of the covariances by $q(t) = \frac{\tau_m^2}{2\tau_s} e^{-|t|/\tau_s}$.

shown in Figure 6, their temporal structure resembles those measured in cortex in vivo [20, Figure 1e,f]: covariances between afferents of the same type are monophasic and positive, while the covariances between excitatory and inhibitory afferents are biphasic and mostly negative. The lag reported between inhibitory and excitatory activity [20, Figure 2b], which was also observed in binary random networks [11, 16], is explained by the echo of the spike contributing to the covariance function. In contrast to previous work, we take the delayed and pulsed synaptic interaction into account. Without delays and with binary neurons [11, 16] the echo appears to be a time lag of inhibition with respect to excitation.

Measuring membrane potential fluctuations during quiet wakefulness in the barrel cortex of mice [31], showed that correlations between inhibitory neurons are typically narrower than those between two excitatory neurons [31, their Figure 4A, 5B and Figure 5C,E]. These results qualitatively agree with our theory for covariances between the spiking activity, because fluctuations of the membrane potential are uniformly transferred to fluctuations of the instantaneous firing intensity. The direct measures of spiking activity [31, their Figure 6] confirm the asymmetric correlation between excitation and inhibition. The low correlation between excitatory neurons reported in that study may partly be due to the unnormalized, firing rate dependent measure and the low rates of excitatory neurons.

Oscillations in the γ range (25 – 100 Hz) are ubiquitous in population measures of neural activity in humans, and have earlier been explained in networks of leaky integrate-and-fire model neurons [9] by the Hopf bifurcation induced by delayed negative feedback. For the regime of high noise here we uncover a simpler analytical condition for the onset (11) and frequency (12) of fast global oscillations. For lower noise, deviations of the non-linear leaky integrate-and-fire dynamics from the linear theory presented here are expected.

From a physics viewpoint, neuronal networks unite several interesting properties. They do not reach thermodynamic equilibrium even in the stationary state, as no detailed balance condition holds between pairs of states of the system. This can be appreciated directly from the fact that a constant firing rate in the network amounts to a continuous flux of the neurons' state variables (voltages) over the threshold. Moreover, the interaction between pairs of neurons is directed, delayed, pulsed, and depends on the flux of the sending neuron's state variable at threshold. In contrast, pairwise interactions frequently studied in physics, like the Coulomb interaction or exchange interaction, can be expressed by a pair potential and are thus symmetric (undirected), instantaneous, and depend directly on the state variables (e.g. spatial coordinates or spins) of the pair of interacting particles. Non-equilibrium systems are at the heart of ubiquitous transport phenomena, like heat or electric conduction. Understanding fluctuations in such a system marks the starting point to infer macroscopic properties by the assertion of the fluctuation-dissipation theorem that connects microscopic fluctuations to macroscopic transport properties. Despite the non-equilibrium dynamics and the non-conservative pairwise interaction, in this manuscript we develop a simple analytical framework merely based on linear perturbation theory that explains time dependent covariance functions of the activity of pairs of integrate-and-fire model neurons in a recurrent random network. Formally our approach resembles the step from the kinetic Ising model near equilibrium to its non-equilibrium counterpart, the network of binary neurons [16]. A difference is the spiking interaction considered in our work, which led us to describe each neuron in terms of the flux over threshold (spike train) rather than by its state variables (membrane voltage and synaptic current). In this respect, we follow the established mean-field approach for spiking neuronal systems [9, 14]. While this mean field approach, however, proceeds by assuming vanishing correlations to obtain the dynamics of the ensemble averaged activity, here we derive and solve the self-consistency equation for the pairwise averaged covariances of the microscopic system.

The typical time scale of covariance functions found here coincides with the time window of biological synaptic plasticity rules [21], so that non-trivial interactions of dynamics and structure are expected. It is our hope that the novel capability to resolve the temporal structure of covariances in spiking networks presented here proves useful as a formal framework to further advance the theory of these correlated non-equilibrium systems and in particular serves as a further stepping stone in the endeavor to understand learning from the interplay of neuronal dynamics and synaptic plasticity.

Acknowledgments

We are thankful to Birgit Kriener, Sonja Grün and George Gerstein for discussions and exploratory simulations in fall 2005 that led to the observation of asymmetric covariance functions. Partially supported by the Helmholtz Alliance on Systems Biology, the Next-Generation Supercomputer Project of MEXT, EU Grant 15879 (FACETS), EU Grant 269921 (BrainScaleS). All network simulations carried out with NEST (<http://www.nest-initiative.org>).

Appendix A. Response kernel of the LIF model

The response kernel h_{ij} needs to be related to the dynamics of the neuron model (1). Here we present an approximation of this kernel which is sufficiently accurate to allow quantitative predictions, but yet simple enough to enable an analytical solution for the correlation structure. If the synaptic time constant is short $\tau_s \ll \tau_m$, the synaptic amplitude J can be thought of as the amplitude of the jump in the membrane potential V caused upon arrival of an incoming impulse. If correlations between incoming spike trains are sufficiently small, the first and second moments of the summed impulses $\tau_m \sum_j J_{ij} s_j(t-d)$ are $\mu_i = \tau_m \sum_j J_{ij} r_j$ and $\sigma_i^2 = \tau_m \sum_j J_{ij}^2 r_j$, respectively, if the inputs' statistics can be approximated by Poisson processes of rate r_j each. For small J and a high total rate, the system of differential equations (1) is hence approximated by a stochastic differential equation driven by a unit variance Gaussian white noise ξ

$$\tau_m \frac{dV_i}{dt} = -V_i + I_{s,i}(t)$$

$$\tau_s \frac{dI_i}{dt} = -I_i + \mu_i + \sigma_i \sqrt{\tau_m} \xi(t).$$

The stationary firing rate in this limit is given by [24]

$$\begin{aligned} r_i^{-1} &= \tau_r + \tau_m \sqrt{\pi} (F(y_\theta) - F(y_r)) \\ f(y) &= e^{y^2} (1 + \operatorname{erf}(y)) \quad F(y) = \int^y f(y) dy \\ \text{with } y_{\theta,r} &= \frac{V_{\theta,r} - \mu_i}{\sigma_i} + \frac{\alpha}{2} \sqrt{\frac{\tau_s}{\tau_m}} \quad \alpha = \sqrt{2} |\zeta(\frac{1}{2})|, \end{aligned} \tag{A.1}$$

with Riemann's zeta function ζ . The rate r_i is the density of action potentials per time. The response of the firing density of the neuron i at time point t with respect to a point-like deflection of the afferent input s_j at time point t' defines the response kernel as the functional derivative

$$\begin{aligned} \left\langle \frac{\delta s_i(t)}{\delta s_j(t')} \right\rangle_{\mathbf{s} \setminus s_j} &= \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \langle s_i(t, \{\mathbf{s}(\tau) + \epsilon \delta(\tau - t') \mathbf{e}_j | \tau < t\}) \\ &\quad - s_i(t, \{\mathbf{s}(\tau) | \tau < t\}) \rangle_{\mathbf{s} \setminus s_j} \\ &\stackrel{\text{def}}{=} h_{ij}(t - t') = w_{ij} h(t - t'). \end{aligned}$$

Here we used the homogeneity, namely the identical input statistics of each neuron i , leading to the same temporal shape $h(t)$ independent of i and the stationarity, so that the kernel only depends on the time difference $t - t'$. We chose $h(t)$ to have unit integral and defined w_{ij} as the integral of the kernel. We determine the temporal integral of the kernel as

$$\begin{aligned} w_{ij} &= \int_{-\infty}^{\infty} h_{ij}(t) dt \\ &= \frac{\partial r_i}{\partial r_j}. \end{aligned} \tag{A.2}$$

The second equality holds because the integral of the impulse response equals the step response [32]. Further, a step in the density s_j corresponds to a step of r_j . Up to linear approximation the effect of the step in the rate r_j on the rate r_i can be expressed by the derivative considering the perturbation of the mean μ_i and the variance σ_i^2 upon change of r_j . Using equation (A.1) we note that by chain rule $\frac{\partial r_i}{\partial r_j} = -r_i^2 \frac{\partial r_i^{-1}}{\partial r_j}$. The latter derivative follows as $\frac{\partial r_i^{-1}}{\partial r_j} = \frac{\partial r_i^{-1}}{\partial y_\theta} \frac{\partial y_\theta}{\partial r_j} + \frac{\partial r_i^{-1}}{\partial y_r} \frac{\partial y_r}{\partial r_j}$. The first derivative in both terms yields $\frac{\partial r_i^{-1}}{\partial y_A} = \sqrt{\pi} \tau_m f(y_A)$ with $y_A \in \{y_\theta, y_r\}$. The second derivative evaluates

with $y_A = \frac{A-\mu_i}{\sigma_i} + \frac{\alpha}{2} \sqrt{\frac{\tau_s}{\tau_m}}$ to $\frac{\partial y_A}{\partial r_i} = -\frac{1}{\sigma_i} \tau_m J_{ij} - \frac{A-\mu_i}{\sigma_i^2} \frac{\tau_m J_{ij}^2}{2\sigma_i} = -\tau_m \frac{J_{ij}}{\sigma_i} (1 + \frac{A-\mu_i}{\sigma_i} \frac{J_{ij}}{2\sigma_i})$. So together we obtain

$$\begin{aligned} w_{ij} &= \alpha J_{ij} + \beta J_{ij}^2 \\ \text{with } \alpha &= \sqrt{\pi} (\tau_m r_i)^2 \frac{1}{\sigma_i} (f(y_\theta) - f(y_r)) \\ \text{and } \beta &= \sqrt{\pi} (\tau_m r_i)^2 \frac{1}{\sigma_i} \left(f(y_\theta) J_{ij} \frac{V_\theta - \mu_i}{2\sigma_i^2} - f(y_r) J_{ij} \frac{V_r - \mu_i}{2\sigma_i^2} \right). \end{aligned} \quad (\text{A.3})$$

Appendix B. Cross spectral matrix in frequency domain

The autocovariance $A(\omega)$ in frequency domain ($F(\omega) = \mathcal{F}[f](\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt$) has two different terms. The first term is a constant r due to the spiking with rate r resulting from the delta peak $r\delta(t)$ in time domain. The second term is the continuous function $a_c(t)$, for example due to refractoriness of the neuron. Further follows from $a(t) = a(-t)$ that $A(\omega) = A(-\omega)$. For $\tau > 0$ the covariance matrix fulfills the linear convolution equation (6). As this equation only holds for the positive half of the time axis, we cannot just apply the Fourier transform to obtain the solution. For negative time lags $\tau < 0$ the covariance matrix is determined by the symmetry $\mathbf{c}(\tau) = \mathbf{c}^T(-\tau)$. Here we closely follow [17] and employ Wiener-Hopf theory [26] to derive an equation for the cross spectral matrix in frequency domain that has the desired symmetry and solves (6) simultaneously. To this end we introduce the auxiliary matrix $\mathbf{b}(\tau) = (h * (\mathbf{M}\mathbf{c}))(\tau) + \mathbf{Q}(h * a)(\tau) - \mathbf{c}(\tau)$ for $-\infty < \tau < \infty$. Obviously, $\mathbf{b}(\tau) = 0$ for $\tau > 0$. Since the defining equation for \mathbf{b} holds on the whole time axis, we may apply the Fourier transform to obtain $\mathbf{B}(\omega) = H(\omega) (\mathbf{M}\mathbf{C}(\omega) + \mathbf{Q}A(\omega)) - \mathbf{C}(\omega)$. Solving for \mathbf{C}

$$\mathbf{C}(\omega) = (\mathbf{1} - \mathbf{M}H(\omega))^{-1} (H(\omega)A(\omega)\mathbf{Q} - \mathbf{B}(\omega)) \quad (\text{B.1})$$

and using the symmetry $\mathbf{C}(\omega) = \mathbf{C}^T(-\omega)$ we obtain the equation

$$\begin{aligned} & (H(\omega)A(\omega)\mathbf{Q} - \mathbf{B}(\omega)) (\mathbf{1} - \mathbf{M}^T H(-\omega)) \\ &= (\mathbf{1} - \mathbf{M}H(\omega)) (H(-\omega)A(-\omega)\mathbf{Q}^T - \mathbf{B}^T(-\omega)). \end{aligned}$$

We observe that $\mathbf{Q}\mathbf{M}^T = \mathbf{M}\mathbf{Q}^T$ is symmetric and with $A(\omega) = A(-\omega)$ the term proportional to $|H(\omega)|^2$ cancels on both sides, remaining with

$$\begin{aligned} & H(\omega)A(\omega)\mathbf{Q} + (\mathbf{1} - \mathbf{M}H(\omega)) \mathbf{B}^T(-\omega) \\ &= H(-\omega)A(\omega)\mathbf{Q}^T + \mathbf{B}(\omega) (\mathbf{1} - \mathbf{M}^T H(-\omega)). \end{aligned} \quad (\text{B.2})$$

We next introduce $D(\omega) = H(\omega)A(\omega)$ which we split into $D(\omega) = D_+(\omega) + D_-(\omega)$, chosen such that $d_+(t)$ (in time domain) vanishes for $t < 0$ and $d_-(t)$ vanishes for $t > 0$. Consequently the Fourier transforms of both terms may have poles in distinct complex half-planes: $D_+(\omega)$ may only have poles in the upper half plane $\Im(\omega) > 0$ and the function vanishes for $\lim_{|\omega| \rightarrow \infty, \Im(\omega) < 0} D_+(\omega) = 0$, following from the definition of the Fourier integral. For $D_-(\omega)$ the half planes are reversed. The analytical properties of $H(\omega)$ are thus similar to those of $D_+(\omega)$, those of $B(\omega)$ are similar to $D_-(\omega)$. We sort the terms in (B.2) such that the left hand side only contains terms that vanish at infinity in the lower half plane $\Im(\omega) < 0$, the right hand side those that vanish in infinity in the upper half plane $\Im(\omega) > 0$

$$\begin{aligned} & D_+(\omega)\mathbf{Q} - D_-(-\omega)\mathbf{Q}^T + (\mathbf{1} - \mathbf{M}H(\omega))\mathbf{B}^T(-\omega) \\ & = D_+(-\omega)\mathbf{Q}^T - D_-(\omega)\mathbf{Q} + \mathbf{B}(\omega)(\mathbf{1} - \mathbf{M}^T H(-\omega)). \end{aligned} \quad (\text{B.3})$$

The left hand side consequently is analytic for $\Im(\omega) < 0$ the right hand side is analytic for $\Im(\omega) > 0$, so (B.3) defines a function that is analytic on the whole complex plane and that vanishes at the border for $|\omega| \rightarrow \infty$. Hence by Liouville's theorem it is 0 and we can solve the right hand side of (B.3) for \mathbf{B}

$$\mathbf{B}(\omega) = (D_-(\omega)\mathbf{Q} - D_+(-\omega)\mathbf{Q}^T)(\mathbf{1} - \mathbf{M}^T H(-\omega))^{-1}.$$

Inserted into (B.1) this yields with the definition $\mathbf{P}(\omega) = (\mathbf{1} - \mathbf{M}H(\omega))^{-1}$

$$\begin{aligned} \mathbf{C}(\omega) &= \mathbf{P}(\omega)(H(\omega)A(\omega)\mathbf{Q} - (D_-(\omega)\mathbf{Q} - D_+(-\omega)\mathbf{Q}^T)\mathbf{P}^T(-\omega)) \\ &= \mathbf{P}(\omega)((D_+(\omega) + D_-(\omega))\mathbf{Q}(\mathbf{1} - \mathbf{M}^T H(-\omega)) \\ &\quad - D_-(\omega)\mathbf{Q} + D_+(-\omega)\mathbf{Q}^T)\mathbf{P}^T(-\omega) \\ &= \mathbf{P}(\omega)(D_+(\omega)\mathbf{Q} + D_+(-\omega)\mathbf{Q}^T - A(\omega)|H(\omega)|^2\mathbf{Q}\mathbf{M}^T)\mathbf{P}^T(-\omega). \end{aligned}$$

The latter expression can be brought to the form

$$\begin{aligned} \mathbf{C}(\omega) &= D_+(\omega)\mathbf{P}(\omega)\mathbf{Q} + D_+(-\omega)\mathbf{Q}^T\mathbf{P}^T(-\omega) \\ &\quad + (D_+(\omega)H(-\omega) + D_+(-\omega)H(\omega) - A(\omega)|H(\omega)|^2) \times \\ &\quad \times \mathbf{P}(\omega)\mathbf{M}\mathbf{Q}^T\mathbf{P}^T(-\omega), \end{aligned} \quad (\text{B.4})$$

which has the advantage that the first two terms have poles in distinct half planes $\Im(\omega) > 0$, and $\Im(\omega) < 0$, respectively. This means these terms only contribute for

positive and negative times, respectively, the last term contributes for positive and negative times.

Appendix C. Spectrum of the propagator

With the Fourier representation $H(\omega) = \frac{e^{-i\omega d}}{1+i\omega\tau_e}$ of the delayed exponential kernel (5) the averaged cross spectrum (7) contains the two functions $U(z)$ and $V(z) = |U(z)|^2$ defined on the complex frequency plane $z = i\omega$. These functions may exhibit poles. The function U has a pole z_k whenever the denominator has a single root $H^{-1}(-iz) - L = 0$ which amounts to the condition $(1 + z_k\tau_e)e^{z_k d} = L$. These complex frequencies can be expressed by the Lambert W function, the solution of $We^W = x$ [30], by

$$\begin{aligned} (1 + z_k\tau_e)e^{z_k d} &= L \\ \left(\frac{d}{\tau_e} + z_k d\right)e^{\frac{d}{\tau_e} + z_k d} &= L \frac{d}{\tau_e} e^{\frac{d}{\tau_e}} \end{aligned} \tag{C.1}$$

as

$$z_k = -\frac{1}{\tau_e} + \frac{1}{d} W_k\left(L \frac{d}{\tau_e} e^{\frac{d}{\tau_e}}\right),$$

leading to (9). The Lambert $W_k(x)$ function has infinitely many branches k [30]. The principal branch has two real solutions, if $x > -e^{-1}$. The remaining branches appear in conjugate pairs. For $x < -e^{-1}$ the principal solutions turn into a complex conjugate pair. This happens at sufficiently strong negative coupling L or long delays d

$$\begin{aligned} L \frac{d}{\tau_e} e^{\frac{d}{\tau_e}} &< -e^{-1} \\ L &< -\frac{\tau_e}{d} e^{-\frac{d}{\tau_e} - 1} \\ \text{or} \quad \frac{d}{\tau_e} &> W_0\left(-\frac{1}{Le}\right). \end{aligned}$$

The principal poles may assume positive real values, leading to oscillations. The condition under which this happens can be derived from (C.1). At the point of transition the pole can be written as $z = i\omega_{\text{crit.}}$; it is a solution to $(1 + i\omega_{\text{crit.}}\tau_e)e^{i\omega_{\text{crit.}}d} =$

L . In order for this equation to be fulfilled, the absolute value and the phase must be identical on both sides. The equation for the absolute value requires $1 + (\omega_{\text{crit.}}\tau_e)^2 = L^2$. This means there are only oscillatory solutions, if the magnitude of the feedback exceeds unity $L < -1$. Since the poles come in conjugate pairs, we can assume w.l.o.g. that $\omega_{\text{crit.}} > 0$. The condition for the absolute value hence reads

$$\omega_{\text{crit.}}\tau_e = \sqrt{L^2 - 1}. \quad (\text{C.2})$$

This is the frequency of oscillation at the onset of the Hopf bifurcation. For strong feedback $|L| \gg 1$ the frequency increases linearly with the magnitude of the feedback. The condition for the agreement of the phases reads $\angle(1 + i\omega_{\text{crit.}}\tau_e) + \omega_{\text{crit.}}d = 0$, so $\frac{\Im(1+i\omega_{\text{crit.}}\tau_e)}{\Re(1+i\omega_{\text{crit.}}\tau_e)} = -\frac{\Im(e^{i\omega_{\text{crit.}}d})}{\Re(e^{i\omega_{\text{crit.}}d})} = \tan \omega_{\text{crit.}}d$, which leads to $\tan \omega_{\text{crit.}}d = -\omega_{\text{crit.}}\tau_e$. This equation has a solution in $\frac{\pi}{2} \leq \omega_{\text{crit.}}d \leq \pi$. In the limit of vanishing delay $d \rightarrow 0$ the frequency goes to infinity, as the solution converges to $\omega_{\text{crit.}}d = \frac{\pi}{2}$. This corresponds to the frequency $f_{\text{crit.}} = \frac{1}{4d}$. Inserting (C.2) leads to $\tan(\frac{d}{\tau_e}\sqrt{L^2 - 1}) = -\sqrt{L^2 - 1}$, which can be solved for the critical delay

$$\frac{d}{\tau_e} = \frac{\pi - \arctan(\sqrt{L^2 - 1})}{\sqrt{L^2 - 1}} \quad (\text{C.3})$$

where we took care that the argument of the tangent is in $[\frac{\pi}{2}, \pi]$. So with (C.2) and (C.3) the oscillatory frequency at the transition can be related to the synaptic delay as

$$2\pi f_{\text{crit.}}d = \omega d = \pi - \arctan(\sqrt{L^2 - 1}).$$

Appendix D. Back transform by residue theorem

In the non-oscillatory state all poles z_k (9) have a negative real part. The function $U(z) = ((1 + z\tau_e)e^{zd} - L)^{-1}$ in (7) then has all poles in the left complex half plane, $\Re(z_k) < 0 \quad \forall k$. We perform the Fourier back transform

$$\begin{aligned} u(t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} U(i\omega) e^{i\omega t} d\omega \\ &= \frac{1}{2\pi i} \oint_{E_t} U(z) e^{zt} dz, \end{aligned} \quad (\text{D.1})$$

replacing the integration path by a closed contour E_t following the imaginary axis from $-i\infty$ to $i\infty$. In order to ensure convergence of the integral, for $t < d$ we need

$\Re(z) > 0$, so we close $E_{t < d}$ in infinity within the right half-plane, where the integrand vanishes. Since there are no poles in the right half-plane, for $t < d$ the path $E_{t < d}$ does not enclose any poles, so $u(t) = 0$. For $t \geq d$ the path $E_{t \geq d}$ must be closed in the left half-plane to ensure convergence of (D.1), so the residue theorem yields

$$u(t) = \Theta(t - d) \sum_{z_k \in E_{t > d}} \text{Res}(U, z_k) e^{z_k t}. \quad (\text{D.2})$$

The residue can be calculated by linearizing the denominator of $U(z_k + z)$ around z_k

$$\begin{aligned} (1 + (z_k + z)\tau_e)e^{(z_k + z)d} - L &= (1 + z_k\tau_e)e^{z_k d}e^{zd} - L + z\tau_e e^{z_k d}e^{zd} \\ &= L(e^{zd} - 1) + z\tau_e e^{z_k d}(1 + zd) + O(z^2) \\ &= z(Ld + \tau_e e^{z_k d}) + O(z^2), \end{aligned}$$

which yields

$$\begin{aligned} \text{Res}(U, z_k) &= \lim_{z \rightarrow 0} zU(z_k + z) \\ &= \lim_{z \rightarrow 0} \frac{z}{z(Ld + \tau_e e^{z_k d})} = \frac{1}{Ld + \tau_e e^{z_k d}} \\ &= \frac{e^{-z_k d}}{(1 + z_k\tau_e)d + \tau_e}, \end{aligned}$$

where we used (C.1) in the last step. The poles of $V(z) = U(z)U(-z)$ are located in both half-planes, consequently $v(t)$ is nonzero on the whole time axis. Here we only calculate $v(t)$ for positive times $t > 0$, because it follows for negative times by symmetry $v(-t) = v(t)$. The path has to be closed in the left half-plane, where the poles z_k have the residues

$$\text{Res}(V, z_k) = \text{Res}(U, z_k)U(-z_k) = \frac{1}{(1 + z_k\tau_e)d + \tau_e} \frac{1}{(1 - z_k\tau_e) - Le^{z_k d}}.$$

So applying (D.2) the functions u and v are

$$\begin{aligned} u(t) &= \sum_{k=0}^{\infty} \frac{1}{(1 + z_k\tau_e)d + \tau_e} \Theta(t - d) e^{z_k(t-d)} \\ v(t) &= \sum_{k=0}^{\infty} \text{Res}(V, z_k) (\Theta(t) e^{z_k t} + \Theta(-t) e^{-z_k t}) \\ &= \sum_{k=0}^{\infty} \frac{1}{(1 + z_k\tau_e)d + \tau_e} \frac{1}{(1 - z_k\tau_e) - Le^{z_k d}} e^{z_k |t|}. \end{aligned} \quad (\text{D.3})$$

The amplitude of the modes decrease with k . For all figures in the manuscript we truncated the series after $k = 30$.

Appendix E. Simulation parameters used for figures

All network simulations were performed using NEST [33]. The parameters of the leaky integrate-and-fire neuron model (1) throughout this work are $\tau_m = 20$ ms, $\tau_s = 2$ ms, $\tau_r = 2$ ms, $V_\theta = 15$ mV, $V_r = 0$ mV. All simulations are performed with precise spike timing and time stepping of 0.1 ms [34]. Figure 1 and Figure 3-Figure 6 of the main text all consider recurrent random networks of N excitatory and γN inhibitory leaky integrate-and-fire model neurons receiving input from randomly drawn neurons in the network and external excitatory and inhibitory Poisson input, so that the first and second moments are $\mu_i = 15$ mV and $\sigma_i^2 = 10$ mV, respectively. Unless stated explicitly, we use $N(1 + \gamma) = 10000$ neurons except in Figure 3 where the number of neurons is given in the legend. Each neuron has $K = pN$ incoming excitatory synapses with synaptic amplitude J independently and randomly drawn from the pool of excitatory neurons, and $\gamma K = \gamma pN$ inhibitory synapses with amplitude $-gJ$ (homogeneous Erdős-Rényi random network with fixed in-degree), realizing a connection probability $p = 0.1$. Cross covariance functions are measured throughout as the covariance between two disjoint populations of 1000 neurons each taken from the indicated populations in the network. Correlation functions are evaluated with a time resolution of 0.1 ms.

In Figure 3 we keep the feedback of the population rate constant $L = Kw(1 - \gamma g) = \text{const.}$ Increasing the size of the network N the synaptic amplitude J (which is proportional to w in linear approximation) needs to scale as $J = J_0 N_0 / N$, where we chose $J_0 = 0.1$ mV and $N_0 = 10000$ here. The variance caused by local input from the network then decreases with increasing network size $\propto 1/N$, while the local mean is constant because $L = \text{const.}$ Each cell receives in addition uncorrelated external balanced Poisson input, adjusted to keep the mean $\mu_i = 15$ mV, and fluctuations $\sigma_i = 10$ mV constant. This is achieved by choosing the rates of the external excitatory ($r_{\text{e,ext.}}$, amplitude $J_{\text{ext.}} = 0.1$ mV) and inhibitory ($r_{\text{i,ext.}}$, amplitude $-gJ_{\text{ext.}}$) inputs as

$$\begin{aligned} r_{\text{e,ext.}} &= r_{\text{e,0}} + r_{\text{bal}} & r_{\text{i,ext.}} &= r_{\text{bal}}/g \\ \text{with } r_{\text{e,0}} &= \frac{\mu_i - \mu_{\text{loc.}}}{J_{\text{ext.}} \tau_m} & \text{and } r_{\text{bal}} &= \frac{\sigma_i^2 - \sigma_{\text{loc.}}^2 - \tau_m r_{\text{e,0}} J_{\text{ext.}}^2}{\tau_m J_{\text{ext.}}^2 (1 + g^2)}, \end{aligned} \tag{E.1}$$

where $\mu_{\text{loc.}} = \tau_m r K J (1 - \gamma g)$ and $\sigma_{\text{loc.}}^2 = \tau_m r K J^2 (1 + \gamma g^2)$ are the mean and variance due to local input from other neurons of the network firing with rate r . From $L = \text{const.}$ follows that also $Kw = \text{const.}$, so that (8) predicts a scaling of the magnitude of the covariance functions in proportion to $1/N$. Other network parameters are $d = 3$ ms and $g = 5$. The firing rate in the network is $r = 23.6$ Hz.

In Figure 4 and Figure 5 we use $N = 10^4$, $g = 6$, $J = 0.1$ mV, and the delay d as described in the captions, the remaining parameters are as in Figure 3.

In Figure 6 we use $N = 10^4$, $J = 0.1$ mV, $g = 5$, and $d = 2$ ms and the remaining parameters as in Figure 3. We obtain the filtered synaptic currents by filtering the spike trains with an exponential filter of time constant $\tau_s = 2$ ms. This results in an effective filter for the cross covariances of $q(t) = \frac{\tau_m^2}{2\tau_s} e^{-|t|/\tau_s}$. The different contributions shown are $c_{I_e I_e} = q * (J^2 p K a_e + J^2 K^2 c_{ee})$, $c_{I_i I_i} = q * (J^2 p K a_i + J^2 K^2 c_{ii})$, $c_{I_e I_i} = -q * (g J^2 \gamma K^2 c_{ei})$, and $c_{I_i I_e} = -q * (g J^2 \gamma K^2 c_{ie})$, where $*$ denotes the convolution.

For Figure 2, we simulate two populations of $N = 1000$ neurons each. Each neuron receives independent background activity from Poisson processes and in addition input from a common Poisson process with rate $r_c = 25$ Hz causing in population 1 a positive synaptic amplitude of J and for population 2 a negative synaptic amplitude J (J is given on the x-axis). The synaptic amplitude of the background inputs is $J_e = 0.1$ mV for an excitatory impulse and $J_i = -0.5$ mV for an inhibitory impulse. The rates of the excitatory and inhibitory background inputs are chosen so that the first and second moments $\mu_i = \tau_m (J_e r_e + J_i r_i + J r_c) = 15$ mV and $\sigma_i^2 = \tau_m (J_e^2 r_e + J_i^2 r_i + J^2 r_c^2) = 10$ mV are independent of J . The spikes produced by each population are triggered to the arrival of an impulse in the common input and averaged over a duration of 10 s to obtain the impulse response.

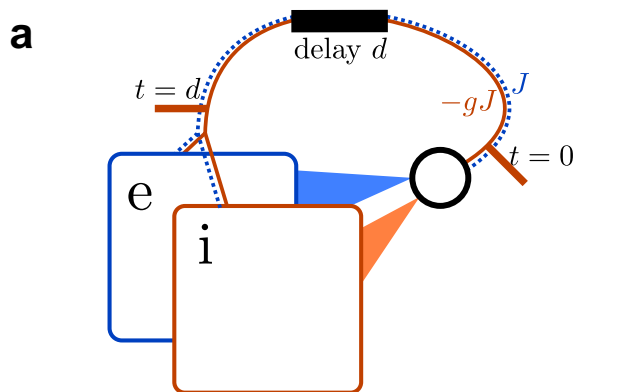
References

- [1] Gerstein, G. L., & Perkel, D. H. (1969). Simultaneously recorded trains of action potentials: analysis and functional interpretation. *Science* 881(164), 828–830.
- [2] Cohen, M. R., & Kohn, A. (2011). Measuring and interpreting neuronal correlations. *Nat. Rev. Neurosci.* 14(7), 811–819. doi:10.1038/nrn.2842.
- [3] Zohary, E., Shadlen, M. N., & Newsome, W. T. (1994). Correlated neuronal

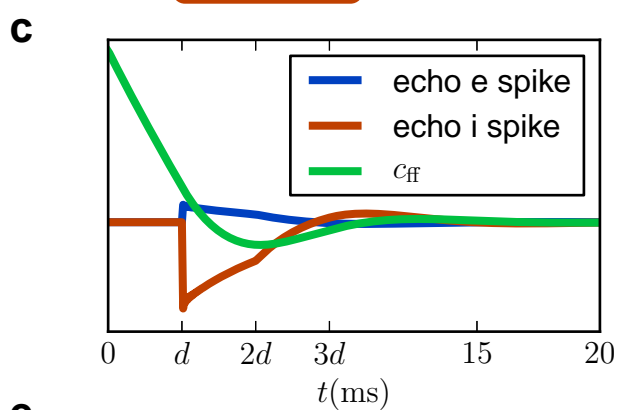
- discharge rate and its implications for psychophysical performance. *Nature* 370, 140–143.
- [4] Abbott, L. F., & Dayan, P. (1999). The effect of correlated variability on the accuracy of a population code. *Neural Comput.* 11, 91–101.
- [5] von der Malsburg, C. (1986). Am I thinking assemblies? In G. Palm & A. Aertsen (Eds.), *Brain Theory*, pp. 161–176. Berlin: Springer-Verlag.
- [6] Maldonado, P., Babul, C., Singer, W., Rodriguez, E., Berger, D., & Grün, S. (2008). Synchronization of neuronal responses in primary visual cortex of monkeys viewing natural images. *100*(3), 1523–1532.
- [7] Kilavik, B. E., Roux, S., Ponce-Alvarez, A., Confais, J., Gruen, S., & Riehle, A. (2009). Long-term modifications in motor cortical dynamics induced by intensive practice. *J. Neurosci.* 29, 12653–12663.
- [8] Ecker, A. S., Berens, P., Keliris, G. A., Bethge, M., & Logothetis, N. K. (2010). Decorrelated neuronal firing in cortical microcircuits. *Science* 327(5965), 584–587.
- [9] Brunel, N. (2000). Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *J. Comput. Neurosci.* 8(3), 183–208.
- [10] Hertz, J. (2010). Cross-correlations in high-conductance states of a model cortical network. *Neural Comput.* 22, 427–447.
- [11] Renart, A., De La Rocha, J., Bartho, P., Hollender, L., Parga, N., Reyes, A., & Harris, K. D. (2010). The asynchronous state in cortical circuits. *Science* 327, 587–590.
- [12] Tetzlaff, T., Helias, M., Einevoll, G. T., & Diesmann, M. (2012). Decorrelation of neural-network activity by inhibitory feedback. *arXiv*, 1204.4393v1 [q-bio.NC].
- [13] van Vreeswijk, C., & Sompolinsky, H. (1996). Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science* 274, 1724–1726.
- [14] Amit, D. J., & Brunel, N. (1997). Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cereb. Cortex* 7, 237–252.
- [15] Kriener, B., Tetzlaff, T., Aertsen, A., Diesmann, M., & Rotter, S. (2008). Correlations and population dynamics in cortical networks. *Neural Comput.* 20, 2185–2226.
- [16] Ginzburg, I., & Sompolinsky, H. (1994). Theory of correlations in stochastic neural networks. *Phys. Rev. E* 50(4), 3171–3191.

- [17] Hawkes, A. (1971). Point spectra of some mutually exciting point process. *J. R. Statist. Soc. Ser. B* 33(3), 438–443.
- [18] Stein, R. B. (1965). A theoretical analysis of neuronal variability. *Biophys. J.* 5, 173–194.
- [19] Rauch, A., La Camera, G., Lüscher, H., Senn, W., & Fusi, S. (2003). Neocortical pyramidal cells respond as integrate-and-fire neurons to in vivo like input currents. *J. Neurophysiol.* 90, 1598–1612.
- [20] Okun, M., & Lampl, I. (2008). Instantaneous correlation of excitation and inhibition during sensory-evoked activities. *Nat. Neurosci.* 11(5), 535–537.
- [21] Morrison, A., Diesmann, M., & Gerstner, W. (2008). Phenomenological models of synaptic plasticity based on spike-timing. *Biol. Cybern.* 98, 459–478.
- [22] Cai, D., Tao, L., Shkarayev, M. S., Rangan, A. V., McLaughlin, D. W., & Kovačič, G. (2012). The role of fluctuations in coarse-grained descriptions of neuronal networks. *Comm. Math.* 10(1), 307–354.
- [23] Cai, D., Tao, L., & McLaughlin, D. W. (2004). An embedded network approach for scale-up of fluctuation-driven systems with preservation of spike information. *Proc. Natl. Acad. Sci. USA* 101(39), 14288–14293.
- [24] Fourcaud, N., & Brunel, N. (2002). Dynamics of the firing probability of noisy integrate-and-fire neurons. *Neural Comput.* 14, 2057–2110.
- [25] Helias, M., Deger, M., Rotter, S., & Diesmann, M. (2010). Instantaneous non-linear processing by pulse-coupled threshold units. *PLoS Comput Biol* 6(9), e1000929. doi:10.1371/journal.pcbi.1000929.
- [26] Hazewinkel, M. (Ed.) (2002). *Encyclopaedia of Mathematics*. Springer.
- [27] Pernice, V., Staude, B., Cardanobile, S., & Rotter, S. (2011). How structure determines correlations in neuronal networks. *PLoS Comput. Biol.* 7(5), e1002059.
- [28] Trousdale, J., Hu, Y., Shea-Brown, E., & Josic, K. (2012). Impact of network structure and cellular response on spike time correlations. *PLoS Comput. Biol.* 8, e1002408.
- [29] Guillozic, S., L’Heureux, I., & Longtin, A. (1999). Small delay approximation of stochastic delay differential equations. *Phys. Rev. E* 59, 3970–3982.
- [30] Corless, R. M., Gonnet, G. H., Hare, D. E. G., Jeffrey, D. J., & Knuth, D. E. (1996). On the lambert w function. *Advances in Computational Mathematics* 5, 329–359.

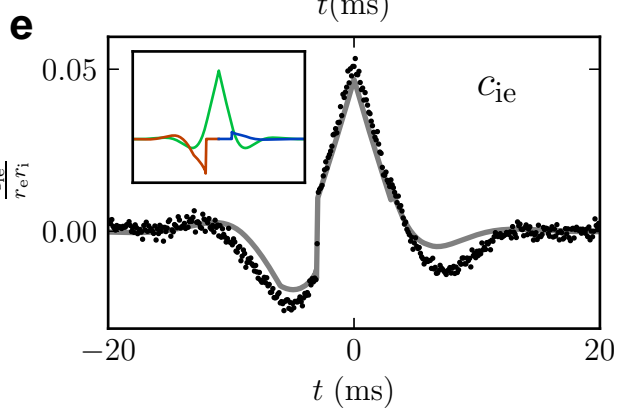
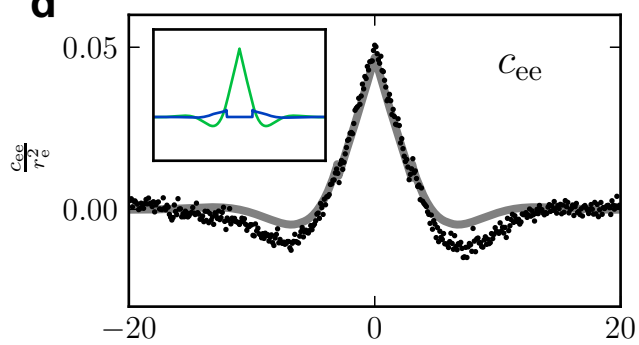
- [31] Gentet, L., Avermann, M., Matyas, F., Staiger, J. F., & Petersen, C. C. (2010). Membrane potential dynamics of GABAergic neurons in the barrel cortex of behaving mice. *Neuron* 65, 422–435.
- [32] Oppenheim, A., & Wilsky, A. (1996). *Systems and signals*. Prentice Hall.
- [33] Gewaltig, M.-O., & Diesmann, M. (2007). NEST (NEural Simulation Tool). *Scholarpedia* 2(4), 1430.
- [34] Hanuschkin, A., Kunkel, S., Helias, M., Morrison, A., & Diesmann, M. (2010). A general and efficient method for incorporating precise spike times in globally time-driven simulations. *Front. Neuroinform.* 4, 113.



b



d



f

